

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.

Application Serial Number: 10/529,931
Source: PCT
Date Processed by STIC: 4-18-05

ENTERED

**CRF Errors Edited by the STIC Systems
Branch**

Serial Number: 10/529,931

CRF Edit Date: 4-18-05
Edited by: GL

☐ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

☐ Corrected the SEQ ID NO. Sequence numbers edited were:

☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

☒ Deleted: ☒ invalid beginning/end-of-file text ; _____ page numbers

☐ Inserted mandatory headings/numeric identifiers, specifically:

☐ Moved responses to same line as heading/numeric identifier, specifically:

☐ Other:

Revised 09/09/2003



PCT

RAW SEQUENCE LISTING

DATE: 04/18/2005

PATENT APPLICATION: US/10/529,931

TIME: 16:39:36

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\04182005\J529931.raw

```

4 <110> APPLICANT: Glaxo Group Limited
6 <120> TITLE OF INVENTION: Vaccine
8 <130> FILE REFERENCE: PG4961
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/529,931
C--> 11 <141> CURRENT FILING DATE: 2005-03-31
15 <150> PRIOR APPLICATION NUMBER: GB 0222953.2
16 <151> PRIOR FILING DATE: 2002-10-03
18 <160> NUMBER OF SEQ ID NOS: 28
20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1107
24 <212> TYPE: DNA
25 <213> ORGANISM: HPV
27 <400> SEQUENCE: 1
28 atggaagcta ttgccaagcg actggaagcc tgccaggagc agctgctgga gctgtacgag 60
29 gaaaacagca cagacctcca caagcacgtg ctgcactgga agtgcacgca ccacgagtc 120
30 gtgctcctgt acaaggccaa gcagatgggg ctgtcccaca tcgggatgca ggctcgtgcc 180
31 ccgctgaagg tgagcgaagc caagggccac aacgctatcg agatgcagat gcacctggag 240
32 agcctgctgc ggaccgaata cagcatggag ccctggactc tccaggagac gtcctacgaa 300
33 atgtggcaga ctctccgaa gcgctgttcc gcaaagcgcg gcaagacagt tgaggtgaaa 360
34 ttcgatgggt gcgcaaacaa cagcatggac tacgtgggtg ggaccgatgt ctacgtgcag 420
35 gacaatgaca cctgggtgaa ggtacatagt atggtggatg ccaagggcat ctattacacc 480
36 tgcgggcagt tcaagacgta ctacgtcaac ttcgtcaagg aagccgaaaa gtatggttcc 540
37 accaagcact gggaggtgtg ttacgggagt actgtgatct gcagccccgc ctccgtgtcg 600
38 tccaccaccc aggaagttag cattccggag agcaccacat acaccccggc ccaaacgagc 660
39 acgctcgtca gcagcagcac caaggaggac gccgtccaga cgcccccccg gaagagggcc 720
40 cgggggggtcc agcagtctcc ctgcaatgcc ctgtgcgttg ctacatcgg ccctgtcgat 780
41 tctgggaacc acaatctcat cacgaacaac cagcaccagc accaaaggcg caacaactct 840
42 aacagctccg caactccaat agtgcagttc cagggggagt ccaactgcct caagtgttcc 900
43 cgctaccgcc tcaacgaccg ccaccgccac ctgttcgact tgatcagttc cacgtggcac 960
44 tgggccagca gcaaggcgcc ccacaaacac gctatcgtga cggtgaccta cgactccgag 1020
45 gagcagaggg agcagttcct ggacgtcgtg aagattcctc cgacaatcag ccacaagctt 1080
46 ggcttcatgt ccctgcacct gctgtga 1107
48 <210> SEQ ID NO: 2
49 <211> LENGTH: 368
50 <212> TYPE: PRT
51 <213> ORGANISM: HPV
53 <400> SEQUENCE: 2
54 Met Glu Ala Ile Ala Lys Arg Leu Asp Ala Cys Gln Glu Gln Leu Leu
55 1 5 10 15
56 Glu Leu Tyr Glu Glu Asn Ser Thr Asp Leu His Lys His Val Leu His
57 20 25 30
58 Trp Lys Cys Met Arg His Glu Ser Val Leu Leu Tyr Lys Ala Lys Gln

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/529,931

DATE: 04/18/2005

TIME: 16:39:36

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\04182005\J529931.raw

```

59          35          40          45
61 Met Gly Leu Ser His Ile Gly Met Gln Val Val Pro Pro Leu Lys Val
62          50          55          60
63 Ser Glu Ala Lys Gly His Asn Ala Ile Glu Met Gln Met His Leu Glu
64 65          70          75          80
65 Ser Leu Leu Arg Thr Glu Tyr Ser Met Glu Pro Trp Thr Leu Gln Glu
66          85          90          95
67 Thr Ser Tyr Glu Met Trp Gln Thr Pro Pro Lys Arg Cys Phe Ala Lys
68          100          105          110
69 Arg Gly Lys Thr Val Glu Val Lys Phe Asp Gly Cys Ala Asn Asn Thr
70          115          120          125
71 Met Asp Tyr Val Val Trp Thr Asp Val Tyr Val Gln Asp Asn Asp Thr
72          130          135          140
73 Trp Val Lys Val His Ser Met Val Asp Ala Lys Gly Ile Tyr Tyr Thr
74 145          150          155          160
75 Cys Gly Gln Phe Lys Thr Tyr Tyr Val Asn Phe Val Lys Glu Ala Glu
76          165          170          175
77 Lys Tyr Gly Ser Thr Lys His Trp Glu Val Cys Tyr Gly Ser Thr Val
78          180          185          190
79 Ile Cys Ser Pro Ala Ser Val Ser Ser Thr Thr Gln Glu Val Ser Ile
80          195          200          205
81 Pro Glu Ser Thr Thr Tyr Thr Pro Ala Gln Thr Ser Thr Leu Val Ser
82          210          215          220
83 Ser Ser Thr Lys Glu Asp Ala Val Gln Thr Pro Pro Arg Lys Arg Ala
84 225          230          235          240
85 Arg Gly Val Gln Gln Ser Pro Cys Asn Ala Leu Cys Val Ala His Ile
86          245          250          255
87 Gly Pro Val Asp Ser Gly Asn His Asn Leu Ile Thr Asn Asn His Asp
88          260          265          270
89 Gln His Gln Arg Arg Asn Asn Ser Asn Ser Ser Ala Thr Pro Ile Val
90          275          280          285
91 Gln Phe Gln Gly Glu Ser Asn Cys Leu Lys Cys Phe Arg Tyr Arg Leu
92          290          295          300
93 Asn Asp Arg His Arg His Leu Phe Asp Leu Ile Ser Ser Thr Trp His
94 305          310          315          320
95 Trp Ala Ser Ser Lys Ala Pro His Lys His Ala Ile Val Thr Val Thr
96          325          330          335
97 Tyr Asp Ser Glu Glu Gln Arg Gln Gln Phe Leu Asp Val Val Lys Ile
98          340          345          350
99 Pro Pro Thr Ile Ser His Lys Leu Gly Phe Met Ser Leu His Leu Leu
100          355          360          365
103 <210> SEQ ID NO: 3
104 <211> LENGTH: 1950
105 <212> TYPE: DNA
106 <213> ORGANISM: HPV
108 <400> SEQUENCE: 3
109 atggcagacg attccggtac tgagaacgaa ggttctggtt gtaccggttg gttcatggtt 60
110 gaagcaatcg ttcagcatcc gactggtacc cagatctccg atgacgaaga cgaagaagtt 120
111 gaagattctg gttacgacat ggttgacttc atcgatgact ccaacatcac tcataactct 180

```

RAW SEQUENCE LISTING

DATE: 04/18/2005

PATENT APPLICATION: US/10/529,931

TIME: 16:39:36

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\04182005\J529931.raw

```

112 ctggaagcac aggctctgtt taaccgccag gaagctgata cccattacgc tactgttcag 240
113 gacctgggag gcaaatatct gggctctccg tacgtttccc cgatcaacac tatcgcagaa 300
114 gcagttgagt ctgaaatctc cccgcgcctg gacgctatca aactgactcg tcagccgaag 360
115 aaggttaaac gtcgtctgtt ccagactcgt gaactgaccg actccggtta cggttatagc 420
116 gaagttgagg ctggcaccgg caccaggtt gaaaaacacg gtgtaccgga aaacggcggc 480
117 gacggtcagg aaaaggacac cggccgcgac atcgagggtg aggaacacac cgaagctgaa 540
118 gctccgacta actctgttcg tgaacacgca ggtactgcgg gtatcctgga actgctgaaa 600
119 tgcaaagacc tgcgcgcggc tctgctgggc aaattcaaag aatgcttcgg cctgtctttc 660
120 attgacctga tccgtccgtt taagtctgac aaaactacct gtctggactg ggttgtagca 720
121 ggcttcggca tccaccactc tatctctgaa gcattccaga aactgatcga gccgctgtct 780
122 ctgtacgcgc acatccagtg gctgactaac gcttggggta tggttctgct ggtactgctg 840
123 cgcttttaaag taaacaaatc tcgttccact gttgctcgta ctctggctac cctgctgaac 900
124 atcccggaga accagatgct gatcgaaacc cggaaaatcc agtctgggtg agctgcaactg 960
125 tactggtttc gtactggcat ctctaaccgt agcactgtta tcgggtgaagc accggaatgg 1020
126 atcactcgtc agaccgttat cgaacacggg ctggcagatt ctgagttcaa actgactgaa 1080
127 atggttcagt gggcatacga caacgacatc tgcgaggaat ctgaaattgc gttcgaatac 1140
128 gctcagcgtg gcgacttcga ctccaacgct cgtgctttcc tgaacagcaa catgcaggct 1200
129 aaatacgtaa aagactgcgc taccatgtgc cgtcactaca aacacgcgga aatgcgtaaa 1260
130 atgtctatca aacagtggat caagcacgcg ggttctaata tcgaaggtag cggtaactgg 1320
131 aaacgatcgt ttcagttcct ggcgcacatc aacatogaat tcatcccggt cctgacccaa 1380
132 ttcaagctgt ggctgcacgg taccocgaaa aaaaactgca tcgctatcgt aggtccaccg 1440
133 gacactgaca agtcttaact ctgtatgtcc ctgatctctt tcctgggcgg cactgtaatc 1500
134 tctcacgtta actcttcctc ccatttctgg ctgcagccac tggtagacgc gaaagtagct 1560
135 ctgctggacg acgcgaccca gccgtgctgg atctacatgg atacttacat gcgcaacctg 1620
136 ctggacggta acccgatgtc tatcgaccgt aaacacaaag cgctgactct gatcaagtgc 1680
137 ccgccgctgc tggtaacttc taacatcgac atcaccaagg aagataaata caagtacctg 1740
138 cataccgctg ttactacctt tactttcccg aaccggttcc cgtttgatcg taacggtaac 1800
139 gctgtttacg aactgtccaa cactaactgg aaatgcttct tcgagcgctc gtcttctctc 1860
140 ctggacatcc aggactctga agatgaagaa gatggttcta actctcaggc tttccggtgt 1920
141 gttccgggta ctgttggtcg tactctgtga 1950
143 <210> SEQ ID NO: 4
144 <211> LENGTH: 649
145 <212> TYPE: PRT
146 <213> ORGANISM: HPV
148 <400> SEQUENCE: 4
149 Met Ala Asp Asp Ser Gly Thr Glu Asn Glu Gly Ser Gly Cys Thr Gly
150 1 5 10 15
151 Trp Phe Met Val Glu Ala Ile Val Gln His Pro Thr Gly Thr Gln Ile
152 20 25 30
153 Ser Asp Asp Glu Asp Glu Glu Val Glu Asp Ser Gly Tyr Asp Met Val
154 35 40 45
155 Asp Phe Ile Asp Asp Ser Asn Ile Thr His Asn Ser Leu Glu Ala Gln
156 50 55 60
157 Ala Leu Phe Asn Arg Gln Glu Ala Asp Thr His Tyr Ala Thr Val Gln
158 65 70 75 80
159 Asp Leu Gly Gly Lys Tyr Leu Gly Ser Pro Tyr Val Ser Pro Ile Asn
160 85 90 95
161 Thr Ile Ala Glu Ala Val Glu Ser Glu Ile Ser Pro Arg Leu Asp Ala
162 100 105 110

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/529,931

DATE: 04/18/2005

TIME: 16:39:36

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\04182005\J529931.raw

```

163 Ile Lys Leu Thr Arg Gln Pro Lys Lys Val Lys Arg Arg Leu Phe Gln
164      115      120      125
165 Thr Arg Glu Leu Thr Asp Ser Gly Tyr Gly Tyr Ser Glu Val Glu Ala
166      130      135      140
167 Gly Thr Gly Thr Gln Val Glu Lys His Gly Val Pro Glu Asn Gly Gly
168 145      150      155      160
169 Asp Gly Gln Glu Lys Asp Thr Gly Arg Asp Ile Glu Gly Glu Glu His
170      165      170      175
171 Thr Glu Ala Glu Ala Pro Thr Asn Ser Val Arg Glu His Ala Gly Thr
172      180      185      190
173 Ala Gly Ile Leu Glu Leu Leu Lys Cys Lys Asp Leu Arg Ala Ala Leu
174      195      200      205
175 Leu Gly Lys Phe Lys Glu Cys Phe Gly Leu Ser Phe Ile Asp Leu Ile
176      210      215      220
177 Arg Pro Phe Lys Ser Asp Lys Thr Thr Cys Leu Asp Trp Val Val Ala
178 225      230      235      240
179 Gly Phe Gly Ile His Ser Ile Ser Glu Ala Phe Gln Lys Leu Ile
180      245      250      255
181 Glu Pro Leu Ser Leu Tyr Ala His Ile Gln Trp Leu Thr Asn Ala Trp
182      260      265      270
183 Gly Met Val Leu Leu Val Leu Leu Arg Phe Lys Val Asn Lys Ser Arg
184      275      280      285
185 Ser Thr Val Ala Arg Thr Leu Ala Thr Leu Leu Asn Ile Pro Glu Asn
186      290      295      300
187 Gln Met Leu Ile Glu Pro Pro Lys Ile Gln Ser Gly Val Ala Ala Leu
188 305      310      315      320
189 Tyr Trp Phe Arg Thr Gly Ile Ser Asn Ala Ser Thr Val Ile Gly Glu
190      325      330      335
191 Ala Pro Glu Trp Ile Thr Arg Gln Thr Val Ile Glu His Gly Leu Ala
192      340      345      350
193 Asp Ser Gln Phe Lys Leu Thr Glu Met Val Gln Trp Ala Tyr Asp Asn
194      355      360      365
195 Asp Ile Cys Glu Glu Ser Glu Ile Ala Phe Glu Tyr Ala Gln Arg Gly
196      370      375      380
197 Asp Phe Asp Ser Asn Ala Arg Ala Phe Leu Asn Ser Asn Met Gln Ala
198 385      390      395      400
199 Lys Tyr Val Lys Asp Cys Ala Thr Met Cys Arg His Tyr Lys His Ala
200      405      410      415
201 Glu Met Arg Lys Met Ser Ile Lys Gln Trp Ile Lys His Arg Gly Ser
202      420      425      430
203 Lys Ile Glu Gly Thr Gly Asn Trp Lys Pro Ile Val Gln Phe Leu Arg
204      435      440      445
205 His Gln Asn Ile Glu Phe Ile Pro Phe Leu Thr Lys Phe Lys Leu Trp
206      450      455      460
207 Leu His Gly Thr Pro Lys Lys Asn Cys Ile Ala Ile Val Gly Pro Pro
208 465      470      475      480
209 Asp Thr Asp Lys Ser Tyr Phe Cys Met Ser Leu Ile Ser Phe Leu Gly
210      485      490      495
211 Gly Thr Val Ile Ser His Val Asn Ser Ser Ser His Phe Trp Leu Gln

```

RAW SEQUENCE LISTING

DATE: 04/18/2005

PATENT APPLICATION: US/10/529,931

TIME: 16:39:36

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\04182005\J529931.raw

```

212          500          505          510
213 Pro Leu Val Asp Ala Lys Val Ala Leu Leu Asp Asp Ala Thr Gln Pro
214          515          520          525
215 Cys Trp Ile Tyr Met Asp Thr Tyr Met Arg Asn Leu Leu Asp Gly Asn
216          530          535          540
217 Pro Met Ser Ile Asp Arg Lys His Lys Ala Leu Thr Leu Ile Lys Cys
218 545          550          555          560
219 Pro Pro Leu Leu Val Thr Ser Asn Ile Asp Ile Thr Lys Glu Asp Lys
220          565          570          575
221 Tyr Lys Tyr Leu His Thr Arg Val Thr Thr Phe Thr Phe Pro Asn Pro
222          580          585          590
223 Phe Pro Phe Asp Arg Asn Gly Asn Ala Val Tyr Glu Leu Ser Asn Thr
224          595          600          605
225 Asn Trp Lys Cys Phe Phe Glu Arg Leu Ser Ser Ser Leu Asp Ile Gln
226          610          615          620
227 Asp Ser Glu Asp Glu Glu Asp Gly Ser Asn Ser Gln Ala Phe Arg Cys
228 625          630          635          640
229 Val Pro Gly Thr Val Val Arg Thr Leu
230          645
233 <210> SEQ ID NO: 5
234 <211> LENGTH: 1104
235 <212> TYPE: DNA
236 <213> ORGANISM: hpv.
238 <400> SEQUENCE: 5
239 atggaagcca tcgcaagag gctcgacgcc tgccaggacc agctgctcga gctgtacgag 60
240 gagaacagca ttgacatcca taagcacatc atgcactgga agtgcattcg cctggagagc 120
241 gtgctgttgc acaaggccaa gcagatgggc ctgtcccaca taggccttca ggtgggtccc 180
242 cctctgaccg tgtcagagac aaagggccat aacgcaatcg agatgcagat gcacctcgag 240
243 tcgctggcga aaacacagta cggcgtggag ccatggaccc tgcaggacac ctctacgaa 300
244 atgtggctga cccacctaa gcgatgttc gccaaacagg gcaacacagt ggaggtgaag 360
245 ttcgacggct gtgaggataa cgttatggag tatgtcgtgt ggacgcacat ctatctgcag 420
246 gacaacgaca gttgggtgaa ggtgaccagc tccgtggacg cgaagggcat ctactatacc 480
247 tgtgggcagt ttaaaaccta ctatgtgaac ttcaacaaag aggcccaaaa gtatggctcc 540
248 accaaccact gggaggtctg ctatgggagc acggtgattt gctctccgc cagcgtgtct 600
249 agcactgtgc gcgaggtgag cattgccgag ccgaccacgt acaccctgc ccagacgacc 660
250 gctccgaccg tgtctgcttg tactaccgag gacggcgtga gcgctccacc caggaagcgt 720
251 gcgagggggc caagcaccaa caacaccctc tgtgtggcga acattcgag cgctcgacagt 780
252 accatcaata acatcgtgac ggataactat aacaagcacc agaggcgtaa caactgtcac 840
253 tctgccgcaa ccccatcgt gcagctccag ggagacagca attgccttaa gtgcttccgc 900
254 tatcgctca acgacaagta caagcacctc tttgagctcg cctcgctgac gtggcactgg 960
255 gcctcaccgg aggcacctca caagaacgcc atcgtcactc tcaacttactc cagtgaggag 1020
256 cagagacagc agtttctgaa cagcgtgaag atcccaccga cgatccgtca taaggtcggc 1080
257 ttcattgtcac tgcattctct gtga 1104
259 <210> SEQ ID NO: 6
260 <211> LENGTH: 367
261 <212> TYPE: PRT
262 <213> ORGANISM: HPV
264 <400> SEQUENCE: 6
265 Met Glu Ala Ile Ala Lys Arg Leu Asp Ala Cys Gln Asp Gln Leu Leu

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/529,931

DATE: 04/18/2005

TIME: 16:39:37

Input Set : A:\pto.kd.txt

Output Set: N:\CRF4\04182005\J529931.raw

L:10 M:270 C: Current Application Number differs, Replaced Current Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date